Name: Matrikelnummer:
Bachelor/Master

Wintersemester 2001/2002 **22. April 2002**
Frank Hoffmann, Christian Knauer

Nachklausur zur Vorlesung
Algorithmen und Datenstrukturen (für Bioinformatik)

1.	2.	3.	4.	Σ
/5	/8	/8	/9	/30

Beginn: $16^{\underline{00}}$, Ende: $18^{\underline{00}}$ (120 min.)

Außer Schreibutensilien sind keine Hilfsmittel erlaubt!

Auf diesem Klausurbogen ist genügend Platz, um die Lösungen der Aufgaben aufzuschreiben. Notfalls kann auch die Rückseite eines Blattes verwendet werden (bitte auf der Vorderseite anmerken). **Zusätzliche Blätter** müssen mit der Matrikelnummer versehen werden. *Der Klausurbogen ist auf jeden Fall mit abzugeben!*

Nicht vergessen, auf allen Blättern die Matrikelnummer einzutragen, auf diesem Deckblatt auch den Namen sowie ihren Studiengang!

Matrikelnummer:

1. O-**Notation** (5 Punkte). /5

Beweisen Sie, daß falls $f(n) = \Theta(h(n))$ und $g(n) = \Theta(h(n))$, auch $f(n) + g(n) = \Theta(h(n))$.

Hinweis zur Erinnerung: $f(n) = \Theta(h(n))$ genau dann wenn $f(n) = O(h(n))$ und $f(n) = \Omega(h(n))$.

Matrikelnummer:

2. **Einfarbige kürzeste Wege** (8 Punkte). /8

Gegeben sei ein gerichteter Graph $G = (V, E)$ in dem jede Kante entweder rot oder blau gefärbt ist. Geben Sie einen Algorithmus an, der in polynomieller Zeit für eine beliebige Eingabe G, $s, t \in V$, und $k \in \mathbb{N}$ entscheidet, ob es in G einen einfarbigen Pfad von s nach t gibt, der höchstens k Kanten benutzt. Was ist die Laufzeit Ihres Verfahrens?

Matrikelnummer:

---

3. **Identität zyklischer Strings** (8 Punkte). /8

Gegeben seien zwei *zyklische* Strings s_1, s_2 gleicher Länge n mit $s_1 = x_1 \dots x_n$ und $s_2 = y_1 \dots y_n$ wobei $x_i, y_j \in \Sigma = \{A, G, T, C\}$ für $1 \leq i, j \leq n$.

(a) Geben Sie einen $O(n)$ Algorithmus an, der entscheidet, ob s_2 eine zyklische Rotation von s_1 ist. Ihr Algorithmus soll alle zyklischen Rotationen t von s_1 um höchstens n Zeichen finden, so daß $x_{(i+t) \bmod n} = y_i$ für alle $1 \leq i \leq n$. (3/8)

(b) Das Komplement $\overline{s}$ eines zyklischen Strings $s \in \Sigma^+$ erhält man, wenn man in s jeweils wechselweise die Buchstaben A durch T sowie G durch C ersetzt (z.B. $\overline{ATGC} = TACG$). Erweitern Sie Ihr Verfahren, so daß Sie damit in $O(n)$ Zeit feststellen können, ob s_2 eine zyklische Rotation von s_1 oder $\overline{s_1}$ ist. (3/8)

(c) Das reverse zyklische String s_1^{rev} von s_1 ist $s_1^{rev} = x_n \dots x_1$. Erweitern Sie Ihr Verfahren, so daß Sie damit in $O(n)$ Zeit feststellen können, ob s_2 eine zyklische Rotation von s_1, $\overline{s_1}$, s_1^{rev}, oder $\overline{s_1}^{rev}$ ist. (2/8)

Matrikelnummer:

4. **Anzahl eindeutiger Teilstrings** (9 Punkte). /9

Gegeben sei ein String s der Länge n.

(a) Geben Sie einen Algorithmus an, der in $O(n)$ Zeit die Anzahl der Teilstrings von s bestimmt, die genau einmal in s vorkommen. (3/9)

(b) Erweitern Sie Ihren Algorithmus so, daß er für ein festes $k \geq 1$ die Anzahl der Teilstrings von s bestimmt, die genau k-mal in s vorkommen. (3/9)

(c) Geben Sie eine Familie von Strings $s_1, s_2, s_3, \ldots$ über einem endlichen Alphabet an, die folgende Eigenschaften hat:

   i. Für alle $n \geq 1$ ist $|s_n| = O(n)$, und

   ii. Die Anzahl der Teilstrings von s_n die genau einmal vorkommen ist $\Omega(n^2)$.

(3/9)

Hinweis: Benutzen Sie Suffixbäume. Beachten Sie, daß Kantenbeschriftungen in Suffixbäumen eine Länge größer Eins haben können.